

Prosodic features of recorded English teaching materials: A pilot study

Margaret Maeda

Abstract

Recorded English teaching materials are almost always advertised as “natural” or sometimes as “authentic” even when they sound quite unnatural and are almost certainly not authentic in the sense of being unscripted conversation. The aim of this study was to find characteristic prosodic features of teaching tapes and compare them with dialogue not originally recorded for second language learners of English. Extracts were taken from materials recorded for beginners to intermediate learners, including radio English lessons and audiocassette tapes, and they were compared with extracts from TV documentaries. The teaching material was made for the Japanese and the international markets. All were recorded by American English speakers. They were judged for naturalness by four native-speaking E.F.L. teachers. Using a Kay Elemetric Computerized Speech Laboratory, the following features were examined: intonation patterns, mean pitch of male and female speakers, speech rates and articulation rates, and duration of syllables. Judgement of the tapes and acoustic analysis suggested that much of the speech recorded for teaching was not similar to speech

recorded for native-speaking listeners. However, American-made teaching tapes were judged as more natural than Japan-made tapes, and acoustic analysis suggested that some features of natural speech were more skillfully imitated.

Background

No instrumental studies of the prosodic features of teaching tapes have been carried out, as far as the writer is aware. Research areas relevant to the present study are foreigner talk, teacher talk and baby talk, which are all forms of simplified speech addressed to language learners. Ferguson (1981: 15) classifies teacher talk as a subtype of foreigner talk, and in a comparison table with baby talk, Hatch (1983a: 155) lists studies of both foreigner and teacher talk together under the heading of foreigner talk. Studies of teacher talk have looked at differences between talk outside and inside the classroom on the levels of discourse, syntax, and lexis, but there has been less study of the phonology of teacher talk, especially at the prosodic level (mainly pitch, duration and loudness) (for summaries of teacher talk studies, see Chaudron, 1988, Chapter 3, and Hatch, 1983a: 155-158). Accounts of the prosody of teacher talk have been criticized as being impressionistic (Chaudron, 1988; Griffiths, 1991) with very little attempt to quantify the features described. They mention "exaggerated intonation," (Hatch, 1983a: 66) "louder" and "slow" speech (Henzl, 1973: 217), and longer pauses (Hakansson, 1986), but except for the last writer, who used an oscillogram to measure pause duration in the classroom talk of

teachers of Swedish as a second language, no instrumental measurements have been made, except for stopwatch measurement of speech rate.

Foreigner or teacher talk and baby talk have many prosodic features in common: exaggerated intonation, high mean pitch, slow speech rate, and longer and more frequent pauses coinciding more with phrase boundaries than adult NS - NS (native speaker to native speaker) speech. Ferguson (1981) and Hatch (1983a: 155) have drawn attention to these similarities. In first language acquisition research, the prosodic features of baby talk have been measured acoustically as part of carefully designed experiments (Garnica, 1977; Fernald & Simon, 1984; Fernald et al, 1989). Garnica (1977) compared the speech of mothers on verbal tasks carried out with three experimental groups: adults, two-year-olds and five-year-olds. She measured mean pitch and pitch range and found significant differences between speech addressed by mothers to adults and two-year-olds, but not between speech addressed to adults and five-year-olds. Speech addressed to two-year-olds had a higher mean pitch by about 70 Hz (197.6 Hz to adults; 267.3 Hz to two-year-olds) and had a wider pitch range. Measurement of the duration of content words showed prolonged duration on two content words in a sentence where the same sentences spoken to adults only had one, so that two words were accented instead of the usual one per sentence or phrase. Features which were found in baby talk which were not found in the present study of teaching tapes, were rising pitch terminals on statements and imperatives, and whispering.

Fernald & Simon (1984) compared speech by German mothers to

adults and to newborns. Some of the features they measured were mean pitch, pitch contour types, and mean F_0 excursion, which is a similar measurement to pitch range. Other features were length of pauses, their correspondence with sentence boundaries, and articulation rate, which is the total number of syllables divided by the total speech time, exclusive of pauses. They found a mean pitch of 257 Hz in speech addressed to babies, compared with 203 Hz in speech addressed to adults, and they found wider pitch excursions. They also found whispering. Another feature not characteristic of teacher talk was repeated pitch patterns, either upward or downward glides, which may have the function of communicating maternal affect. In a later study; Fernald et al (1989) studied the speech of both mothers and father to infants of around one year. The parents were native speakers of Italian, German, French, British English, American English and Japanese. All groups of parents raised their mean pitch when talking to infants and most groups widened their pitch range, but fathers did so less than mothers. American English speakers used the most exaggerated prosody when speaking to infants. Women raised their mean pitch from 206 Hz to 308 Hz, and men from 105 Hz to 146 Hz. The tasks in these experiments were similar to teaching in the classroom: telling stories, giving instructions, or talking about pictures.

Conditions for recording teaching tapes are different from classroom teaching in some respects. Speakers on tapes cannot adjust to the level of the listeners by interacting with them, but only to the estimated level of the learners. It consists of instructions, explanations, example sentences, like speech in the classroom, but the use of dialogues on tapes

makes acting ability more important than it is for classroom teaching. The dialogues on teaching tapes are intended to simulate natural conversation, and nearly all textbooks carry claims that accompanying taped material is "natural."

Abercrombie (1965: 2) divides spoken language into three categories: reading aloud, monologue and conversation. Conversation is by far the most commonly occurring form of spoken language. The only way of obtaining recordings of genuine spontaneous conversation is to record surreptitiously (p. 5). If speakers know they are being recorded the conversation is not spontaneous. Reading aloud forms a much smaller proportion of all spoken language than conversation, but it is more common than most people suppose. For example, presenters of documentaries may not be speaking with a script, but the text of their speech starts out in written form. Reading aloud or speaking from a memorized script is an adaptation of conversation. Abercrombie calls it "spoken prose" (p. 4). Spoken prose has "phonetic peculiarities" (p. 2) which distinguishes it from real conversation, among them: (1) Evenness of tempo compared with the frequent variations in tempo of conversation. (2) Pauses closely related to the grammatical structure of the sentences. In real conversation, the end of a sentence "is more likely to be shown by the intonation than by a pause, and pauses often come between two words in close grammatical connection" (p. 8).

Abercrombie (1965: 9) states that "when we, as language teachers, claim we are teaching the spoken language, most of the time what we are teaching is spoken prose". According to Abercrombie's definitions of real conversation and spoken prose, dialogues on teaching tapes

should be counted as spoken prose, not real conversation, because the dialogues have been written before being recorded.

Although tapes are described as natural even in beginning textbooks, it is doubtful that it is possible to use real conversation as listening material with low level learners. Abercrombie (1965: 6) describes the content of real conversation as "pretty horrifying... sometimes unintelligible... and disorganized." The tempo of natural conversation is variable, with speakers sometimes reaching double average speeds. In a paper on speech rates in British English, Tauroza and Allison (1990) state that "teachers and researchers consider the speed of speech to be one of the key factors affecting listening comprehension" (p. 90). In other words, if the speech is too fast, learners will not understand it. Hatch (1983b: 81), not writing about listening materials, asserts, "As teachers we believe that a pedagogical sequence from simple to complex is important for language teaching." But in commenting on a study of speech rates of teachers (1983a: 81), she says, "Teacher 10 must be one of the few teachers in the world who has really internalized our teacher-training directions on using a natural rate of speech when teaching beginners." This contradiction is reflected in the claims, even on beginning textbooks, that recordings on tapes are natural, even when they are far from being natural.

Griffiths (1990) investigated the effect of different speech rates on the comprehension of low intermediate second language learners. His hypothesis was that average speech rates would give higher scores than moderately fast rates, and slow rates, higher scores than average rates. Fast speech resulted in reduced comprehension but he found that slow

rates did not result in higher scores than average rates. He found that lower intermediate students could cope with natural speaking rates if the language was carefully controlled, and few benefits would be obtained from speaking slowly. Kelch (1985) compared the effect of input modified for speech rate and syntactic complexity. He found comprehension to be little affected by syntactic modification, but much improved by slower speech rate.

In a carefully controlled experiment, Tauroza and Allison (1990) investigated speech rates in speech addressed to native listeners in a variety of situations, in order to provide standards to assess the speed of commercially produced listening materials for second language learning. The varieties of speech recorded were radio news broadcasts and documentaries, unscripted conversations, and prepared interviews. 95 percent of their data fell within a range of 190 to 320 syllables per minute, and based on the results of their study, they produced the following table of speech rates:

Table 1.

Estimate of standard rates of speech (syllables per minute) according to Tauroza & Allison 1990.

Fast	=	above 320
Moderately fast	=	280-320
Average	=	230-280
Moderately slow	=	190-230
Slow	=	below 190

Method

Materials

Broadly, two kinds of recordings were used in this study: extracts from recorded teaching material, and extracts from documentaries made by American TV news stations. There were two types of teaching material: made in Japan for the Japanese market (seven recordings), and made in the U.S. for the international market (six recordings).

The teaching materials were thirteen 20 – 30 second extracts from dialogues used with course material for learners from beginning to intermediate level. Two were radio English lessons for Japanese junior high school students and two were tapes used with course books in Japanese junior high schools. The three false beginner tapes made in Japan are aimed at university students who have had six years of English in high school. The materials made for the international market were taken from widely used, popular textbooks. The dialogues chosen were intended for use as listening practice rather than for presenting new grammatical structures, because oral presentation of new structures is often deliberately slow and clear. Some of the material was designed for use with the listeners looking at scripts, particularly lower level material, and some was not (see Appendix C). Dialogues were selected which had one male and one female speaker.

The extracts from the documentaries were one extract each from CNN and CBS programs. The CNN and CBS recordings were not originally made for teaching purposes, but the CNN tape used in this

study is sold as English-teaching material, and other CBS recordings are also sold as teaching material for advanced students. The CNN extract is an interview of a well-known American male actor about his family life. The CBS recording is of three female members of the public who were the victims of scams. The CNN and CBS tapes are not genuine conversation as Abercrombie defined it, as they were not secretly recorded, and although they were almost certainly not scripted, they were probably prepared. Obtaining samples of genuine conversation by secret recording would present ethical difficulties, but this material can reasonably be called authentic listening material in the sense that it is designed for native speakers to listen to rather than language learners. The CNN and CBS tapes were used as samples of near natural speech for acoustic comparison with the teaching tapes, and as a control in the evaluation of the teaching tapes to check if the

Table 2.

Materials examined

Level of textbook	Type of material	No. of extracts
Beginner	J jhs radio series, Levels 1 & 2	2
	J jhs textbook, Level 1 & 3	2
	International	1
False beginner	Japanese	3
	International	2
Intermediate	International	3
Authentic	CNN CBS	2

Key:

International: Textbook produced in the U.S. for the international market.

Japanese: Textbook produced for the Japanese market.

J jhs Japanese junior high school

judges would give high marks to material which was not made for teaching purposes.

Evaluation of naturalness of tapes

The fifteen extracts were evaluated for naturalness by four native speaking teachers of English as Foreign Language. They were asked to make a quick judgement of the naturalness of each recording according to a six-point scale, taking into account the level of the tape (for the evaluation sheet used, see Appendix A). The level of each tape was indicated on the evaluation sheet as Beginner, False Beginner, Intermediate or Advanced (CNN and CBS). These were broader categories than indicated on the textbooks. For the exact levels indicated on the textbooks, see Appendix B.

Using a Kay Elemetric Computerized Speech Laboratory, the mean pitch of speakers, speech rates, articulation rates, and the duration of syllables were measured. Intonation patterns were evaluated visually, because a suitable method of measurement had not been decided on for this study. The teaching tapes were compared with each other and with the CNN and CBS recordings.

Since this is a pilot study, using a small number of judges and an insufficiently large amount of data on which to run reliable statistical tests, the reporting of results is descriptive and tentative. A larger study is being prepared which will use more judges and more taped materials.

Results

Mean pitch

Mean pitch differs depending on a number of factors including discourse and situational factors, individual differences and group accents (Graddol, 1986). Fernald et al (1989) found that mean pitches of British subjects talking to adults were higher than Americans' mean pitches (British / American men: 127 / 105 Hz; British / American women: 222 / 206 Hz). Garnica's (1977) figures for American mothers talking to adults were similar to Fernald et al's figures: 198 Hz and 203 Hz. Fant (1956, cited in Laver, 1994: 451) gives 120 Hz for men and 220 Hz for women in conversational speech in European languages. Graddol (1986) has 219 Hz for British men and 122 Hz for British women reading dialogues, which is the same kind of reading as on the teaching tapes, except that Graddol's study was not related to language acquisition. Fant's maximum ranges are 50 – 250 Hz for men and 120 – 480 Hz for women (1956, cited in Laver, 1994).

Figure 1a) shows that the mean pitches of men on the tapes are high even compared with Fant's 120 Hz, which is higher than Fernald's 105 Hz. It also shows that there is a trend downward in pitch with increasing level of tape. More data will be needed to confirm this trend. Figure 1b) shows that the mean pitches of women on the tapes are high, but the trend downwards is less clear. Women's range is wider than men's so there is more variability in the data.

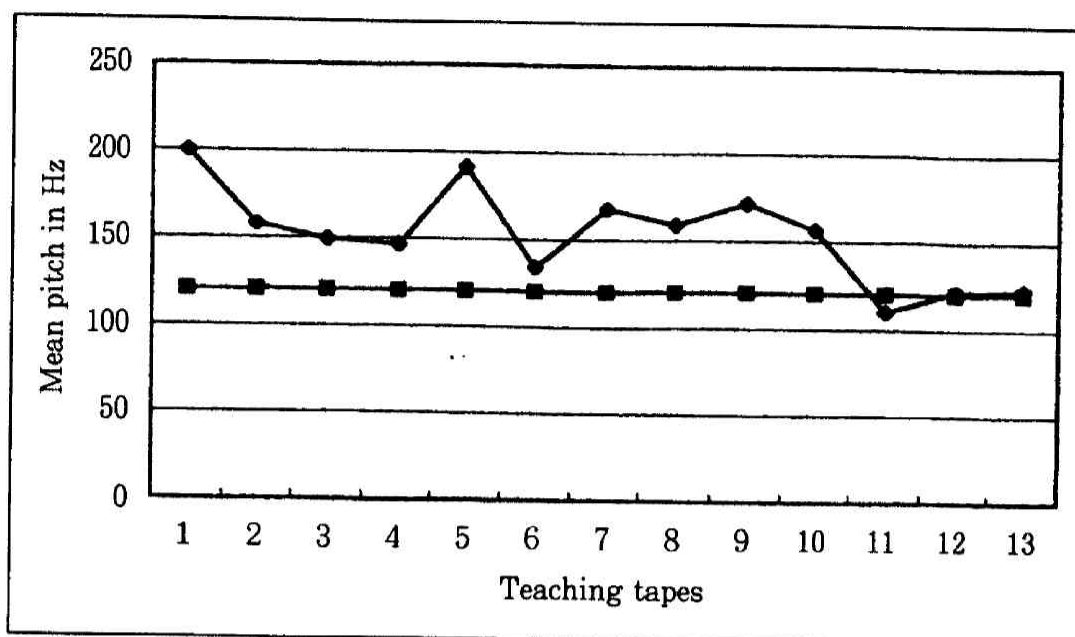


Figure 1a) Mean pitch of men

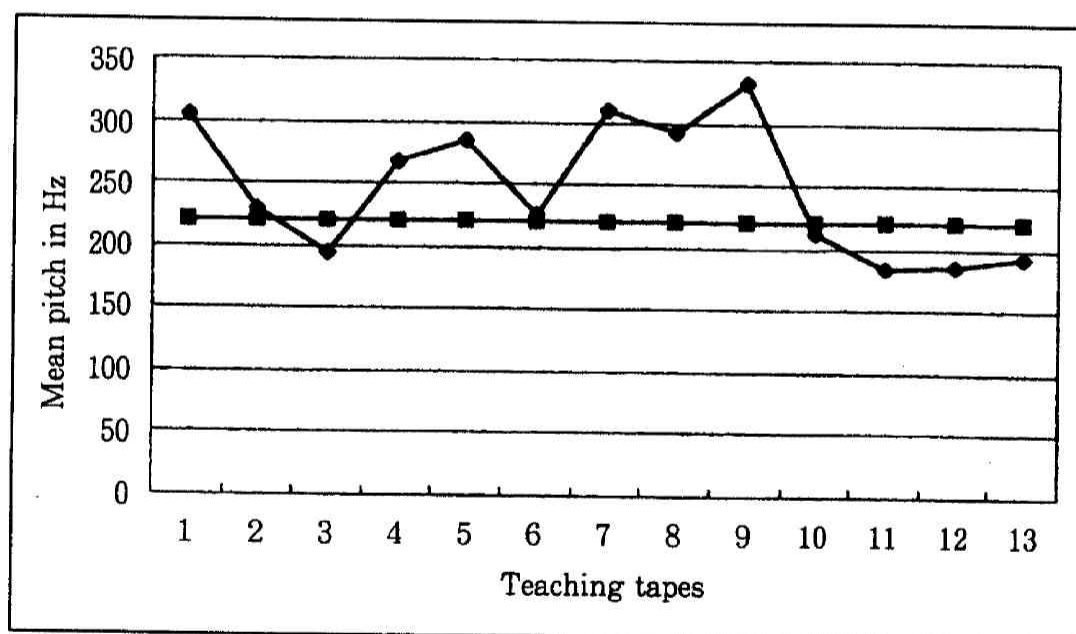


Figure 1b) Mean pitch of women

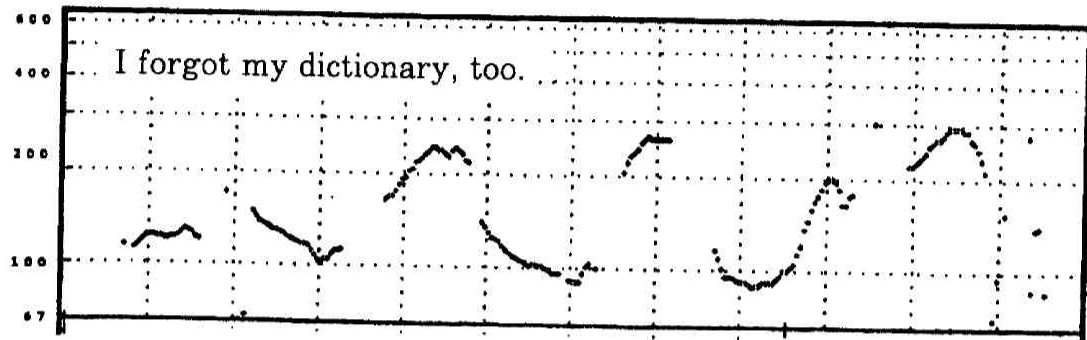
Figure 1. Mean pitch of men and women on teaching tapes compared with conversational speech (Fant 1956: 120 Hz for men, 220 Hz for women).

The mean pitch of men and women on the low level tapes made for the Japanese market was high, and the pitch of the intermediate tapes and the low level tapes made for the international market was near average (see Appendix B). Some of the pitch peaks of both the men and the women, over 500 Hz for women and as high as 300 Hz for men, were extremely high. All the men who reached 300 Hz sounded as if they were straining their voices. However, one of the tapes made in Japan (Tape number 9 in Appendixes B, C, and D), which was very high pitched, received quite favorable evaluations despite sounding unnatural to the writer. One explanation may be that intonation is so variable, depending on discourse and situational factors and individual differences, that a variety of different pitch ranges is acceptable to listeners. The judges may have responded more to another prosodic feature, syllable length, when evaluating the tapes (see the section "Duration of unstressed syllables" below.)

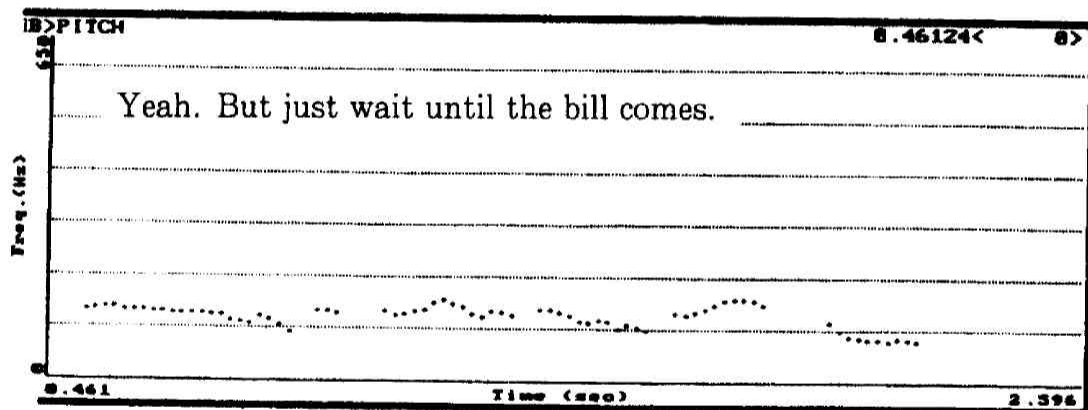
Pitch Patterns

One of the most striking features of teaching tapes is exaggerated intonation. Some judges remarked on "over-acting" in the tapes. Low level tapes often had several wide pitch excursions in a phrase, with a pitch peak on every content word, giving sentences an up-down rhythm. A sentence taken from a beginning teaching tape made in Japan has five words with three large pitch peaks. (see Figure 2a). Fernald & Simon (1989) found more accented words per phrase in baby talk than in adult-to-adult speech. In first language acquisition it is believed that making words stand out helps infants to acquire language. It is likely that for

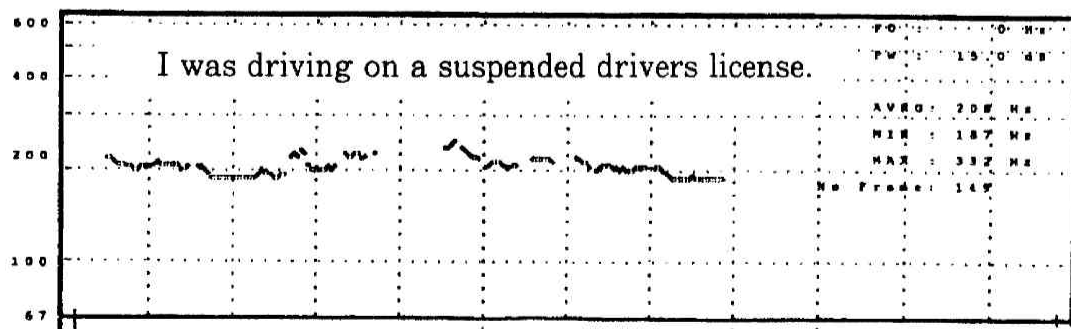
second language learners, accenting words makes them easier to perceive. In spontaneous speech there is, at most, one pitch prominence per phrase, usually the last content word, and the rest of the phrase is



2a) Beginners' tape. (Man)



2 b) Intermediate tape. (Man)



2 c) CBS (woman)

Figure 2. Pitch patterns on teaching tapes.

typically quite flat. The intonation contours of the intermediate tapes (see Figure 2b) were similar to those on the CBS and CNN tapes. A sentence from the CBS tape (Figure 2c) has eight words with one small pitch peak on "suspended," and the phrase has a narrow pitch range.

Baby talk studies give figures for pitch range, but there are problems in measuring pitch range to show exaggerated intonation. Two sentences of the same length, one with one pitch peak and the other with several, could have the same pitch range. Graddol (1986) discusses problems with measuring pitch range and the various methods researchers have used to measure it.

Speech rates

Speech rates (see Appendix C) increased approximately with the level of the materials as advertised by the publishers, and reached rates at the low end of average for speech to native speakers (Tauroza & Allison 1990, see Table 1) at the intermediate level. Griffiths (1990) found that low intermediate students could cope with listening passages which were spoken at average rates. The fairly close relation between increase in speech rate and level of material suggests that speech rate on teaching tapes is quite well controlled. However, Griffiths' (1990) finding that very slow rates did not lead to better comprehension than average rates might indicate that the very slowest tapes do not need to be so slow, even for beginners.

It was thought possible that on the faster low level tapes, high speech rates might be compensated for by less syntactic complexity and shorter utterance length, but calculation of average number of words per T-unit

(a syntactic main clause and its associated subordinate clauses) did not show any consistent inverse relation between length of T-unit and speech rate. A larger sample of tapes and more extracts per tape will need to be analyzed to confirm or refute this hypothesis.

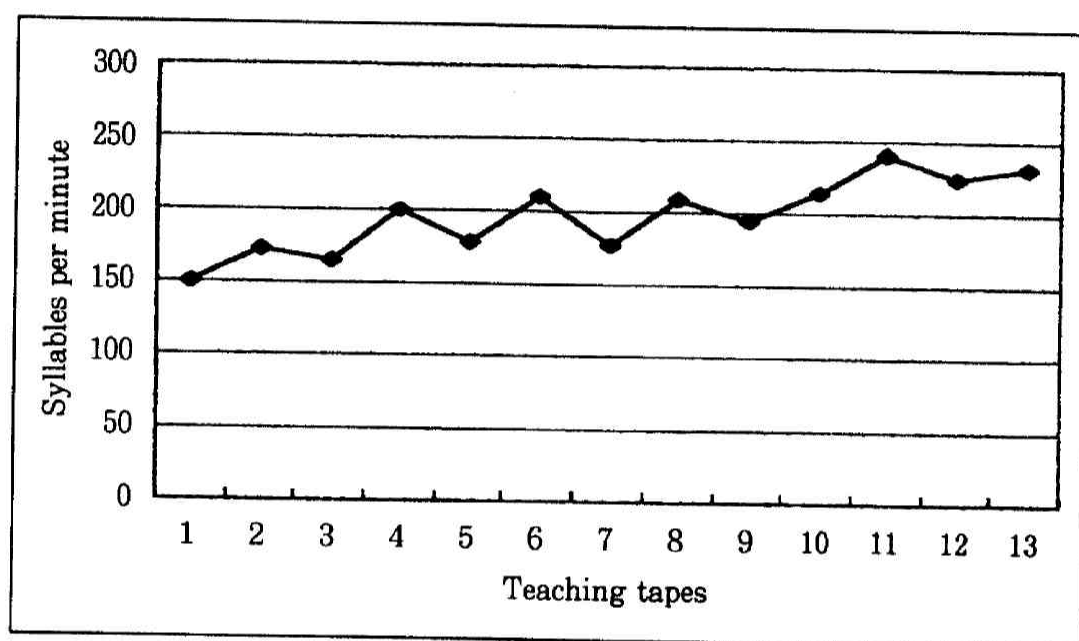


Figure 3. Speech rates in teaching materials.

Articulation rates

The rate of articulation is the speech rate excluding pause time. Pauses at the end of phrases and sentences are known as boundary or juncture pauses, and those within phrases are classified as hesitation pauses. Speed of speaking is more dependent on frequency and duration of pauses than it is on speed of articulation (Goldman-Eisler, 1968), and there are physical limits to the speed at which sounds can be articulated. Goldman- Eisler gives figures cited from Miller (1951) of a maximum of between 6.7 and 8.2 syllables per second and an average of 5 to 5.5 sps. She found a range of 4.4 to 5.9 syllables per second

among the eight subjects she used in an experiment, and a five times greater range of pause time than articulation time.

Fernald & Simon (1984) found an average of 5.8 sps for German mothers speaking to adults and 4.2 sps for the same mothers speaking to babies. In this study, the range was 3.84 to 5.02 sps for the teaching tapes (see Appendix D). Griffiths (1991) minimizes the role of slower enunciation in teacher talk. He cites Goldman-Eisler's claim that speech rate is largely determined by the duration and frequency of pauses and that native speakers rarely draw out their syllables, but her research was not connected with either first or second language acquisition.

Duration of unstressed syllables

In collecting data for this study, there was a striking scarcity of short

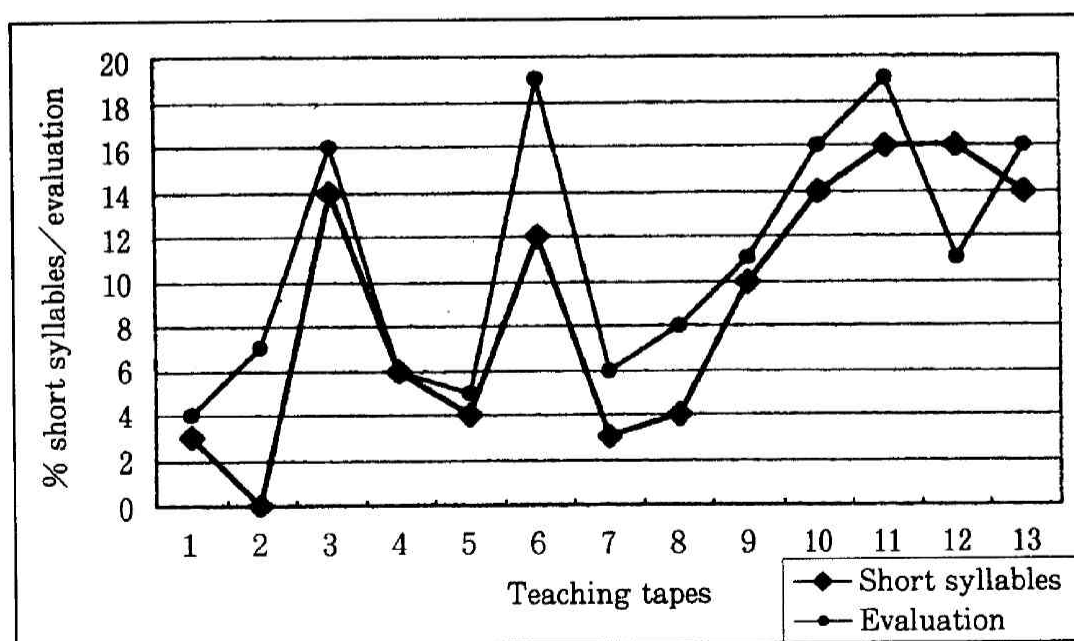


Figure 4. Percentages of short syllables per extract and evaluation of extracts.

syllables in the slow-sounding extracts. In NS-NS speech and some teaching tapes, unstressed syllables and unstressed words like "a" or "the" can be less than 40 ms (milliseconds) long, and are commonly less than 80 ms. In some slow extracts, the absence or near absence of syllables under 100 ms was conspicuous. Low evaluations given by the judges to tapes were remarkably consistent with low percentages of short syllables (see Appendix D and Figure 5). Low evaluation was more consistent with lack of short syllables than mean pitch or speech rate.

Summary

Chaudron (1988) reports that a number of teacher talk studies (including Henzl, 1973, and Hakansson, 1986) show that teachers' speech rates increase in relation to the level of the learners. The same seems to be true of teaching tapes even though the speakers are not interacting with learners. The tapes which had a low percentage of short syllables consistently received low ratings for naturalness. This was the feature most closely tied with low evaluation for naturalness. Only one of the tapes made in Japan did not have a low percentage of short syllables. Interestingly, one of the authors of the textbook is a well-known researcher into second language listening.

Mean pitches were very high compared with NS-NS speech. They decreased with increasing level, but the trend was not so clear as for increasing speech rate. It was less clear for women, who have a wider pitch range, than for men. Pitch fluctuations were wide and numerous

in utterances on low level tapes, but intonation patterns on intermediate tapes were similar to those in the documentaries, with a narrow pitch range fewer and smaller pitch excursions.

The more natural tapes, mostly the tapes produced for the international market, sounded as if they were recorded by professional actors, and the tapes made in Japan were probably recorded by language teachers. The least natural recordings, particularly those made for junior high school students, gave the impression of being designed to reinforce the teaching of grammatical structures rather than to teach the spoken language. Many excellent textbooks are being produced to teach false beginners in Japan at the university level. They are skillfully designed to engage often unmotivated students by using material which interests this group of students. The tapes which accompany the books have well designed listening tasks, but unfortunately, the acting skill of the speakers is not good. Internationally produced textbooks are not so well designed to appeal to Japanese learners, but the acting of the speakers on the tapes is better. The very slow tapes, with drawn out unstressed syllables and unstressed words like "a" and "the," may be giving listeners too much help and may not be preparing them to move towards authentic listening.

Further Research

Further work will use more native-speaking teachers of American English as judges and will need more intermediate material produced in

Japan and complete beginners material made in the U.S. for a balance between the two types of material. Such a balance would probably show more clearly the differences between listening material recorded by teachers and by actors.

Study of pitch patterns will require examination of more material than can be included in an evaluation of naturalness, because pitch phenomena are more variable than duration-related features like syllable length, articulation rates and speech rates.

Some research has already been undertaken to assess the effects of different speech rates on second language comprehension. Research into the effects of manipulation of the other prosodic features considered in this study, like pausing, syllable duration and pitch patterns, might show what modifications of prosody are helpful or unnecessary in the preparation of graded listening materials.

References

- Abercrombie, D. (1965). Conversation and spoken prose. In *Studies in Phonetics and Linguistics*, Oxford University Press, Oxford.
- Chaudron, C. (1988). *Second Language Classrooms*. Cambridge University Press.
- Cruttenden, A. (1986). *Intonation*. Cambridge University Press.
- Fant, G. (1956). On the predictability of formant levels and spectrum envelopes from formant frequencies. In Halle, M., Lunt, H., & MacLean, H., (eds.). *For Roman Jakobson*, The Hague: Mouton, 109-20.
- Fernald, A. and Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, 20, 104-13.
- Fernald, A., (1989). Taeschner, T., Dunn, J., Papousek, M., Boysson-Bardies, B., Fukui, I. A cross-language study of prosodic modifications in mothers' and

- fathers' speech to preverbal infants. *Journal of Child Language*, 16, 477-501.
- Garnica, O. (1977). Some prosodic and paralinguistic features of speech to young children. In Snow, C. E. and Ferguson, C. E. (eds.), *Talking to Children: Language input and acquisition*. Cambridge, MA: University Press, 63-88.
- Goldman-Eisler, (1968). *Psycholinguistics: Experiments in Spontaneous Speech*. London: Academic Press.
- Graddol, D. (1986). Discourse specific pitch behaviour. In Johns-Lewis, C. (ed.) *Intonation in Discourse*. London: Groom Helm.
- Griffiths, R. T. (1990). Speech rate and NNS comprehension: A preliminary study in time-benefit analysis. *Language Learning*, 40: 311-36.
- Griffiths, R. T. (1991). Pausological research in an L2 context: A rationale, and review of selected studies. *Applied Linguistics*, 12: 345-364.
- Hakansson, G. (1986). Quantitative aspects of teacher talk. In G. Kasper, (ed.) *Learning, teaching and communication in the foreign language classroom*, 83-98. Aarhus Denmark: Aarhus University Press.
- Hatch, E. (1983a). *Psycholinguistics: a Second Language Perspective*. Rowley, Mass.: Newbury House.
- Hatch, E. (1983b). Simplified input and second language acquisition. In Andersen (ed.) *Pidginization and Creolization as Second Language Acquisition*. Rowley, Mass.: Newbury House.
- Henzl, V. (1973). Linguistic register of foreign language instruction. *Language Learning*, 23: 207-222.
- Kelch, K. (1982). Modified input as an aid to comprehension. *Studies in Second Language Acquisition*, 7, 81-90.
- Laver, J. (1994). *Principles of Phonetics*. Cambridge University Press.
- Miller, G. A. (1951). *Language and Communication*. New York: McGraw-Hill.
(Cited in Goldman-Eisler, 1968).
- Tauroza, S., & Allison, D. (1990). Speech rates in British English. *Applied Linguistics*, 11, 90-105.

Appendix A

Teaching Tape Evaluation Sheet with Results Entered

Name of judge:

Date of collection:

Approximate length of teaching experience:

No	Level of material	Grades					
		1. Very poor	2. Poor	3.	4.	5. Good	6. Very good
1	Beginner (J jhs radio)	ABD			C		
2	Intermediate (I)		CD	B	A		
3	False beginner (J)	AD	BC				
4	False beginner (I)				C	ABD	
5	Beginner (J jhs)	ABCD					
6	Advanced (CNN)					C	ABD
7	Intermediate (I)				C	ABD	
8	Beginner (J jhs)	ABD	C				
9	Beginner (I)			BC		AD	
10	False beginner (J)	AC		BD			
11	False beginner (I)		C		A	BD	
12	False beginner (J)		D	ABC			
13	Intermediate (I)				ABCD		
14	Advanced (CBS)					C	ABD
15	Beginner (jhs)	ABD		C			

Key:

Bracketed information was not available to the judges.

ABCD Codes for the four judges.

J Material produced for the Japanese market.

I Material produced international market

jhs Material produced for junior high school students.

Appendix B
Mean Pitch of Male and Female Speakers

	Level of text	Market	Mean pitch of women	Mean pitch of men
1	1 st year jhs	J	305	200
2	1 st year jhs (radio)	J	229	158
3	Beginner	I	194	149
4	2 nd year jhs (radio)	J	267	146
5	3 rd year jhs	J	286	191
6	False beginner	I	226	134
7	False beginner	J	309	168
8	False beginner	J	293	159
9	False beg. - low int.	J	333	171
10	False beg. - low int	I	211	156
11	High beg. - high int.	I	183	110
12	Low int. - int.	I	186	121
13	Intermediate	I	193	122
14	CNN			103,142
15	CBS		146,255,222	

Appendix C

Speech Rates, Mean Number of Words per T-Unit, and Evaluations of Extracts

	Level of text	Market	Syllables per minute	Mean no. of words per T-unit	Evaluation	Text available or not
1	1 st year jhs	J	150	5.0	4	text
2	1 st year jhs (radio)	J	171	3.8	7	text
3	Beginner	I	164	5.0	16	text
4	2 nd year jhs (radio)	J	200	5.6	6	text
5	3 rd year jhs	J	177	4.0	5	text
6	False beginner	I	210	4.3	19	text
7	False beginner	J	176	4.0	6	no text
8	False beginner	J	208	5.8	8	no text
9	False beg. - low int.	J	193	5.1	11	no text
10	False beg. - low int.	I	215	4.8	16	text
11	High beg. - high int.	I	241	5.2	19	text
12	Low int. - int.	I	224	6.4	11	no text
13	Intermediate	I	232	5.7	16	no text
14	CNN: whole interview interviewee		251 273	10.3	23	
15	CBS: 3 speakers		276 272 291	8.8	23	
					2 nd Speaker	

Key:

jhs Material produced for junior high school students.

beg. Beginner

int. Intermediate

Appendix D

Percentages of Short Syllables and Articulation Rates

	Level of text	Market	Percentage of short syllables	Evaluation	AR (syllables per second)
1	1 st year jhs	J	3	4	3.84
2	1 st year jhs (radio)	J	0	7	3.89
3	Beginner	I	14	16	4.24
4	2 nd year jhs (radio)	J	6	6	4.07
5	3 rd year jhs	J	4	5	3.53
6	False beginner	I	12	19	4.48
7	False beginner	J	3	6	4.09
8	False beginner	J	4	8	4.46
9	False beg. – low int.	J	10	11	4.54
10	False beg. – low int.	I	14	16	4.86
11	High beg. – high int.	I	16	19	5.02
12	Low int. – int.	I	16	11	4.59
13	Intermediate	I	14	16	4.92
14	CNN		26	23	6.20
15	CBS		12	23	5.00

Key:

AR: Articulation Rate

英語録音教材の韻律的特徴

前田マーガレット

現在市販されている殆ど全ての日本で制作された英語の録音教材は、不自然な発音の会話が録音されているにも関わらず、誠しやかに“自然”もしくは“本物”などという宣伝文句と共に販売されている。故に筆者はこれら録音テープの韻律的特徴を調べ、またその特徴（ラジオの英会話番組を含む）とアメリカで全世界向けに制作された英語の録音教材とアメリカで放映されたテレビ ドキュメント番組の自然な会話の韻律を比べることとした。方法として、まず先に、初級者と中級者用に作成されたテープ（ラジオ放送教材を含む）とテレビ ドキュメント番組のピッチ曲線をコンピュータ音声解析装置を用い抽出した。これら全てのテープは英語母語話者（アメリカ人）によって録音されたものである。次に、これらの録音テープの一部を再度一本のテープにランダムに録音し、それを4人のE.F.L. 教員（英語母語話者）に聞いてもらい、発音がどれだけ自然であるかランク付けしてもらった。この研究で今回音響的調査した内容は、イントネーション パターン、男女のミーン ピッチ、スピーチ レート、調音度、とシラブルの長さである。この研究でわかったことは殆ど全ての日本国内で作成された教材用テープの発話はテレビ ドキュメント番組内の発話とはまったくと言ってもいい程似ていない、しかしアメリカで制作された教材用テープの発話は日本製の物に比べてはるかに自然で音響的分析においても、これらの発話はナチュラル スピーチに似るように巧妙に発音されている。